Experiments:

1. Normal algorithm with variance in the environment

How does the original algorithm do when changing the variance?
→ compare on each environment with different reward variance means, with rew_var_var = None
(can do 1 graph for each a.1, b.1, b.2, c.1, c.2)

a. Sparse environment
   1. rew_var_mean_ter = None
      rew_var_mean_ter = 100 (hypothesis: a bit slower to converge)
      rew_var_mean_ter = 10 000 (hypothesis: slower to converge)

b. Semi sparse environment
   1. rew_var_mean_ter = None, rew_var_mean_step = None (hypothesis: no difference)
      rew_var_mean_ter = None, rew_var_mean_step = 1 (hypothesis: slower to converge)
   2. rew_var_mean_ter = 10 000, rew_var_mean_step = None
      (hypothesis: slower to converge)
      rew_var_mean_ter = 10 000, rew_var_mean_step = 1
      (hypothesis: very slow to converge)

c. Dense environment
   1. rew_var_mean_ter = None, rew_var_mean_step = None
      rew_var_mean_ter = None, rew_var_mean_step = 1
      (hypothesis: a bit slower to get to max)
      rew_var_mean_ter = None, rew_var_mean_step = 100
      (hypothesis: slower to converge)
   2. rew_var_mean_ter = 10 000, rew_var_mean_step = None
      (hypothesis: slower to converge)
      rew_var_mean_ter = 10 000, rew_var_mean_step = 1
      (hypothesis: slower to converge)
      rew_var_mean_ter = 10 000, rew_var_mean_step = 100
      (hypothesis: really slower to converge)

2. Modified algorithm vs normal algorithms

How does the modified algorithm compare to the normal algorithm?
→ compare on each environment with different reward variance means, and different variance variances:

(need to do one graph for each case a.1, a.2, …, b.1, b.2, … with a curve for each algo + curve for normal algo on env without variance on the reward – to compare how close do we get)

a. Sparse environment
   1. rew_var_mean_ter = None
      (hypothesis: no difference)
   2. rew_var_mean_ter = 10 000, rew_var_var_ter = None
      (hypothesis: no difference)
   3. rew_var_mean_ter = 10 000, rew_var_var_ter = 250 000 (+- 500)
      (hypothesis: modified is a bit better)
   4. rew_var_mean_ter = 10 000, rew_var_var_ter = 25 000 000 (+- 5 000)
      (hypothesis: modified is really better)

b. Semi sparse environment

1. rew_var_mean_ter = None, rew_var_mean_step = None
   (hypothesis: no difference)
2. rew_var_mean_ter = None, rew_var_mean_step = 1, rew_var_var_step = None
   (hypothesis: no difference)
3. rew_var_mean_ter = None, rew_var_mean_step = 1, rew_var_var_step = 0.6
   (hypothesis: modified is better)
4. rew_var_mean_ter = 10 000, rew_var_var_ter = 25 000 000,
   rew_var_mean_step = None
   (hypothesis: modified is better)
5. rew_var_mean_ter = 10 000, rew_var_var_ter = 25 000 000, rew_var_mean_step = 1,
   rew_var_var_step = 0.6
   (hypothesis: modified is way better)

c. Dense environment

1. rew_var_mean_ter = None, rew_var_mean_step = None
   (hypothesis: no difference)
2. rew_var_mean_ter = None, rew_var_mean_step = 1, rew_var_var_step = None
   (hypothesis: no difference)
3. rew_var_mean_ter = None, rew_var_mean_step = 1, rew_var_var_step = 0.6
   (hypothesis: modified is a bit better)
4. rew_var_mean_ter = None, rew_var_mean_step = 100, rew_var_var_step = 6 000
   (hypothesis: modified is a lot better)
5. rew_var_mean_ter = 10 000, rew_var_var_ter = None,
   rew_var_mean_step = None
   (hypothesis: no difference)
6. rew_var_mean_ter = 10 000, rew_var_var_ter = 25 000 000,
   rew_var_mean_step = None
   (hypothesis: modified is better)
7. rew_var_mean_ter = 10 000, rew_var_var_ter = 25 000 000,
   rew_var_mean_step = 1, rew_var_var_step = 0.6
   (hypothesis: modified is better)
7. rew_var_mean_ter = 10 000, rew_var_var_ter = 25 000 000,
   rew_var_mean_step = 100, rew_var_var_step = 6 000
   (hypothesis: modified is way better)